Exploring the operation of SEA in the ACT

Analysis brief

Prepared for the

Australian Capital Territory Education Directorate

October 2021

VICTORIA UNIVERSITY
MELBOURNE AUSTRALIA

CIRES
Centre for International Research on Education Systems

About CIRES

The Centre for International Research on Education Systems, located at Victoria University, conducts strategic research that identifies how well education systems work, for whom, and how they can be improved to work well for all. The Centre undertakes large-scale survey and policy-related projects covering every state and territory in Australia and every sector of education and training. It also undertakes international comparative research examining the features and performance of education systems around the world.

Executive summary

Publicly released comparisons of school performance are high stakes. Parents can use the information when choosing a school for their child, which can also be used by school leaders and others. There is an onus on systems to ensure reported performance is accurate and fair.

The current approach to reporting school performance in NAPLAN seeks to achieve the goals of accuracy and fairness using what is called the similar-schools model. This model compares the original NAPLAN performance of schools, to that expected given the context of students and schools. A key measure of student and school context is student family background, as measured by the occupation and education of parents and carers. Student family background is included in the similar-schools model by a variable called the Socio-Educational Advantage (SEA) index.

The sensitivity and accuracy of school performance measurement using the new modelling approach depends on the quality of data used to measure and adjust for student background and school context. Questions have been raised about the quality of some of the input data used to measure student family background, and the effects of the data in generating potentially perverse results for groups of schools at a system level, such as for the Australian Capital Territory (ACT). The anomalies raise questions about the accuracy and fairness of the school comparisons.

For this reason, the ACT Education Directorate commissioned the Centre for International Research on Education Systems (CIRES) at Victoria University to participate in a project exploring the operation of the Socio-Educational Advantage (SEA) index in the assessment of student and school performance in the ACT.

Key findings from the first phase of analysis

- Analysis was undertaken to compare the education and occupation data reported to ACARA, to that recorded in other sources, including Census of Population and Housing (2016), and the Longitudinal Survey of Australian Children (LSAC). No material inconsistency or anomalies were identified across the three sources.

- The project then examined the effect of the similar-schools model on the reported performance of ACT schools, focusing on Year 9 Numeracy in 2019. When using the original (unadjusted) NAPLAN data, 35.1 per cent of ACT schools are top performers (in the top quartile of schools in Australia). However, the similar-schools modelling approach sees the percentage in the top quartile fall to 2.7 per cent.

- This marked shift in performance is attributable to the similar-schools modelling approach, and particularly the role of the SEA variables. Of the six variables used in the similar-schools model, only the two SEA variables play a role for the vast majority of ACT students when estimating expected performance.

- The results for ACT schools are unchanged by using alternative constructions of SEA, or by how SEA is applied in the similar-schools model. That does not mean that there

are no anomalies in SEA and its construction. Rather, further work on how SEA is constructed, and analysis of the way SEA affects the ACT, may be needed.

- In terms of the similar-schools modelling, there is some evidence that missing variables may be important to study further. Selective school status, when included in the Year 9 Numeracy model, both improves the model explanation of variance, and the results for ACT schools.

- There is no evidence to suggest there is a systemic performance issue in the ACT, given outcomes on other valued student outcome measures. For that reason, further analysis is required to investigate alternative approaches to measuring school performance.

Options for next steps

Based upon the analysis findings to date, there are three avenues that could be further examined.

The first area of additional analysis involves closer analysis of high SEA schools in the ACT and other jurisdictions, examining performance across a range of different measures. The analysis would initially focus upon schools that are performing strongly in the unadjusted NAPLAN scores, but not in the modelled scores. Comparison would be made with other schools also performing strongly in modelled data.

The second area of analysis could look at ways of refining the existing similar-schools modelling. There is potential to explore whether the truncated nature of NAPLAN scores—with a floor and ceiling—is adversely impacting upon the ACT. There is also an opportunity to use data such as the Longitudinal Survey of Australian Children, to identify whether variables representing the 'home learning environment' can further explain student performance than SEA itself. Finally, the role of school selectiveness could be further explored, taking into account the abilities of school cohorts as identified in previous NAPLAN years.

The third area of analysis would look at different ways of measuring school performance, using both NAPLAN data and other datasets. Initially, this might mean undertaking comparisons with approaches to school performance assessments in other countries. Then there is value in examining other school performance measures. Two promising options involving NAPLAN data include (1) using NAPLAN proficiency standards to identify the numbers of students achieving agreed skill levels, and (2) using measures of change in skill levels or learning gain over time employing appropriate adjusted or relative measures of learning gain . Analysis of students achieving a set proficiency standard could use estimates derived from both unadjusted and modelled performance. Adjusted or relative learning gain would compare the gain of students between years 3 and 5, and 7 and 9, adjusting the amount of gain relative to the starting point. This could also be estimated using appropriate performance models to adjust for school contextual differences.

Table of Contents

Executive summary ii
1. Measuring school performance accurately and fairly 1
Main issue 2
Approach to analysing the issue 2
2. Analysis findings 4
Effects of modelling on ACT school performance ratings 4
Between-school differences and SEA 12
Alternative constructions of SEA don't change the results 15
Sensitivity of the 'similar-schools' modelling to SEA construction 16
Alternative applications of SEA in the modelling 17
Relevance of other variables and approaches to ACT results 18
3. Summary of findings and options for next steps 22
Summary 22
Options for next steps 22
References 25

List of Figures

Figure 2-1 Position change of ACT schools following modelling: Year 9 Numeracy 2019 6
Figure 2-2 Position change of New South Wales and ACT schools: Year 9 Numeracy 2019 7
Figure 2-3 Position change of Victorian and ACT schools: Year 9 Numeracy 2019 7
Figure 2-4 Position change of ACT schools following modelling: Year 9 Reading 2019 8
Figure 2-5 Change in percentile rank of New South Wales and ACT schools, by school SEA: Year 9 Reading 2019 9
Figure 2-6 Change in percentile rank of Victorian and ACT schools, by school SEA: Year 9 Reading 2019 9
Figure 2-7 Contribution of student and school characteristics to the modelled results: Year 9 Numeracy 2019 11
Figure 2-8 Between-school variance in Year 9 Numeracy and Reading:2019 (%) 14
Figure 2-9 School-level residual (standardised) and selective schools: ACT and New South Wales schools: Year 9 Numeracy (2019) 21

List of Tables

Table 2-1 Quartile position of schools: percentage of schools, by jurisdiction 5
Table 2-2 Quartile position of students for Year 9 Numeracy: Percentage of students, by jurisdiction 10
Table 2-3 Variance in Year 9 Numeracy and Reading (2019) and SEA_AV quintile 13
Table 2-4 SEA construction approach options 16
Table 2-5 Variance explained and performance classification of ACT schools: SEA variables for 2019 Year 9 Numeracy PVs (%) 17
Table 2-6 Alternative applications of SEA in modelling 18
Table 2-7 Variance explained and performance classification of ACT schools: SEA application for 2019 Year 9 Numeracy PVs (%) 18
Table 2-8 Other variables and approaches 19
Table 2-9 Variance explained and performance classification of ACT schools using additional variables and approaches: 2019 Year 9 Numeracy PVs (%) 20

Acronyms and abbreviations

ACARA	Australian Curriculum, Assessment and Reporting Authority
ICSEA	Index of Community Socio-Educational Advantage
LSAC	Longitudinal Survey of Australian Children
PV	Plausible Value
SEA	Socio-Educational Advantage index
WLE	Weighted Likelihood Estimate

1. Measuring school performance accurately and fairly

Publicly released comparisons of school performance are high stakes. Parents can use the information in choosing a school for their child, school leadership teams can use the information to evaluate how well the school is doing, and school councils can use the information to see if the plans for the school are meeting the expectations of the local community. There is an onus on systems, therefore, to ensure that the assessment of performance is both accurate and fair. In pursuit of accuracy and fairness, Australian schools have seen some innovations in performance comparisons of National Assessment Program—Literacy and Numeracy (NAPLAN) results over recent years. There has been a move from the earlier days of comparing schools based on 'like-school' groupings through to comparing schools based on the modelled results of the performance of all students across the nation, adjusted for context differences associated with such characteristics as parental/carer education and occupation, remoteness of the school and percentage of Indigenous students.

The need to measure performance accurately and fairly by adjusting for school and student differences in context is well accepted. Studies in Australia and around the world tell us that student background (linked to characteristics such as parental/carer education, occupation, poverty, and cultural background) can make a big difference to student learning and progress. So, to judge each school's performance fairly there is a need to take account of differences in the backgrounds and types of students at each school, and differences in the community contexts of schools. An accepted method for doing this is through multi-level modelling such as that now being used in Australia. This method is designed to take into account differences in the context of each school, the challenges faced, and consequently the value being added in improving student learning.

However, the sensitivity and accuracy of school performance measurement using the new modelling approach depends on the quality of data used to measure and adjust for student background and school context. Questions have been raised about the quality of some of the input data used to measure student family background, and the effects of the data in generating potentially perverse results for particular groups of schools, at a system level, such as for the Australian Capital Territory (ACT). The anomalies raise questions about the accuracy and fairness of the school comparisons.

For this reason, the ACT Education Directorate commissioned the Centre for International Research on Education Systems (CIRES) at Victoria University to participate in a project exploring the operation of the Socio-Educational Advantage (SEA) index in the assessment of student and school performance in the ACT. The SEA index is a key variable used in the similar-schools modelling undertaken to classify NAPLAN performance.

The project was undertaken in cooperation with the Australian Curriculum, Assessment and Reporting Authority (ACARA). The terms of reference for the project, agreed between the ACT Education Directorate and ACARA, were to:

assess the potential that an anomaly in the operation of the Socio-Educational Advantage (SEA) index is materially impacting inter-jurisdictional NAPLAN comparisons involving the ACT where these comparisons rely on data that comprise some elements of the SEA.

Main issue

The main concern is that the similar-schools modelling may not be accurately reflecting the performance of schools in the ACT and the SEA variable may be the main reason. In 2019, the Education Council endorsed a revised methodology for benchmarking the performance of Australian schools. Rather than comparing every school against 60 'like-schools' grouped by ICSEA, NAPLAN performance was now measured using a multi-level model (MLM) controlling for a range of student and school characteristics including SEA (as captured by parental education and occupation), remoteness of the school and percentage of Indigenous students. Relative performance of a school is assessed using the difference (residual) between the actual and the predicted school mean performance based on the model results.

Questions have been raised about the new approach to measuring and reporting school performance because of the potential for data anomalies to exist that bias comparisons of schools. The main question relates to the use of the background education and occupation data from parents/carers, as currently collected and subsequently analysed to form the SEA and ICSEA. This is distinct from whether or not the data, and its treatment, results in any systematic difference in the way school performance is predicted for schools in the ACT compared to schools in other parts of Australia.

The main goal of the project is to examine whether the background data from parents/carers, as currently collected and subsequently analysed to form the SEA and ICSEA scales, results in any systematic difference in the way school performance is predicted for schools in the ACT, compared to schools in the rest of the country. If SEA is determined to not be an accurate measure in the ACT due to the presence of anomalies in its operation, the project would also ascertain what effect any anomalies may have on comparative measures of school performance.

There has been little public discussion regarding the use of multi-level modelling to compare Australian school NAPLAN performance. One exception is analysis by the Catholic Education Commission of Victoria, which argued that student cultural/language background should be included as explanatory variables in the modelling (Catholic Education Commission of Victoria, 2019).[1]

Approach to analysing the issue

Several strands of work were undertaken to examine the issue, and are reported in this analysis brief:

[1] The data made available by ACARA for this project did not identify student cultural/language background. Rather, only Language Background Other than English (LBOTE) status was made available.

1. compare the results for ACT schools in performance assessments based on the similar-schools modelling against unadjusted performance to identify the extent to which the similar-schools modelling affects the reported NAPLAN performance of ACT schools
2. explore alternative methods of constructing SEA to that used by ACARA, and examine how the resulting SEA values affect the ACT similar-schools modelling results
3. examine whether or not the inclusion of other variables in the similar-schools modelling affect the performance assessments of ACT schools, and
4. outline options for further analysis.

The analysis reported in this document builds upon previous work undertaken for the current project that examined whether there was a potential inaccuracy in how the background education and occupation of ACT parents is reported. The previous analysis compared the data on parent education and occupation reported to ACARA, to that recorded in both the Australian Bureau of Statistics Census of Population and Housing (2016), and the Longitudinal Survey of Australian Children (LSAC). The previous analysis, undertaken at a jurisdiction level and limited to the parents of school children, did not identify any material differences between the ACARA data, and what is recorded in the both the Census and LSAC.

The earlier analysis also examined whether there were significant differences between the income levels reported in the ABS Census among parents/carers working in the different occupations, between the ACT and other states and territories. Finally, analysis using the Longitudinal Surveys of Australia Youth (2015) examined if there were differences between measures of family wealth and home educational resources, between the ACT and the rest of Australia.

2. Analysis findings

Effects of modelling on ACT school performance ratings

Instead of reporting school performance using the original NAPLAN scores, ACARA uses similar-schools modelling to estimate a student's expected NAPLAN score given information about student and school characteristics (ACARA, 2020). Using resulting school-level averages, the differences between the original and modelled scores provide the basis for the My School reporting.[2] A school is considered to be performing well if its original score is significantly higher than its modelled score.

School level analysis

The model-based approach to measuring NAPLAN school performance adopted by ACARA has a profound effect on the apparent performance of ACT schools. To display this, Table 2-1 compares the quartile position of schools in each state and territory, using the original (unadjusted) NAPLAN data and then the modelled (adjusted) NAPLAN data for Year 5 and Year 9 Numeracy and Reading. All schools were grouped into quartiles based upon the average NAPLAN score of their students. Quartile 1 represent the lowest 25 percent of schools based on NAPLAN numeracy performance. Quartile 4 is the highest 25 per cent performing schools in Australia.

Focussing upon Year 9 Numeracy, the analysis shows that when using the original (unadjusted) NAPLAN data, 35.1 per cent of ACT schools are top performers (in the top quartile of schools in Australia). This is the highest percentage across the 8 states and territories. However, the results of the similar-schools modelling approach sees the percentage in the top quartile fall to just 2.7 per cent. Further, the proportion of ACT schools in the bottom two quartiles increases from just 35.1 per cent (well below any other jurisdiction) to 91.9 per cent (well above any other jurisdiction). Similar patterns were identified in Year 9 Reading, alongside Year 5 Reading and Numeracy. However, the movement between the original (unadjusted) NAPLAN scores, and the modelled results was not as great.

[2] The difference between the school-level original and modelled score is known as the school-level residual. This is also known as the random effect.

Table 2-1 Quartile position of schools: percentage of schools, by jurisdiction

	Quartile 1		Quartile 2		Quartile 3		Quartile 4	
	Original	Modelled	Original	Modelled	Original	Modelled	Original	Modelled
Year 5 Numeracy								
ACT	13.4	82.5	26.8	13.4	43.3	4.1	16.5	0.0
NSW	25.6	18.5	22.5	25.3	23.0	27.1	28.9	29.1
NT	72.4	29.5	16.2	38.1	6.7	24.8	4.8	7.6
QLD	27.5	24.9	27.1	23.0	26.1	26.5	19.3	25.6
SA	38.3	52.3	29.0	28.8	22.4	14.6	10.4	4.4
TAS	38.3	23.5	36.2	27.6	17.4	31.1	8.2	17.9
VIC	11.7	16.8	24.4	22.4	29.9	27.2	33.9	33.6
WA	29.6	33.0	25.0	29.1	22.3	20.8	23.1	17.2
Year 5 Reading								
ACT	10.3	53.6	20.6	30.9	33.0	12.4	36.1	3.1
NSW	26.5	22.9	23.8	24.5	23.6	24.8	26.2	27.8
NT	68.9	53.4	16.5	24.3	9.7	10.7	4.9	11.7
QLD	28.2	22.9	26.3	22.2	24.6	27.6	21.0	27.3
SA	32.4	39.3	28.8	30.2	23.5	22.6	15.3	7.8
TAS	31.1	10.7	28.1	21.9	25.5	23.0	15.3	44.4
VIC	14.1	20.2	24.9	24.5	29.4	27.2	31.6	28.2
WA	28.0	30.0	24.7	28.2	22.4	22.7	24.9	19.1
Year 9 Numeracy								
ACT	21.6	78.4	13.5	13.5	29.7	5.4	35.1	2.7
NSW	26.2	20.9	21.6	23.9	23.4	25.8	28.8	29.5
NT	76.7	32.6	11.6	34.9	7.0	23.3	4.7	9.3
QLD	28.7	30.9	27.2	27.6	26.0	25.2	18.1	16.3
SA	28.6	41.7	33.2	30.2	24.6	21.1	13.6	7.0
TAS	36.5	15.3	37.7	30.6	12.9	23.5	12.9	30.6
VIC	15.4	24.8	26.5	21.7	28.9	25.7	29.2	27.8
WA	20.5	9.3	21.6	24.5	26.4	27.1	31.6	39.0
Year 9 Reading								
ACT	18.9	62.2	10.8	13.5	27.0	13.5	43.2	10.8
NSW	24.8	22.5	22.9	21.9	24.4	24.8	27.9	30.9
NT	69.8	34.9	14.0	37.2	11.6	14.0	4.7	14.0
QLD	31.8	29.5	23.9	27.8	25.6	25.4	18.8	17.3
SA	28.1	35.7	31.7	34.7	24.6	22.1	15.6	7.5
TAS	39.3	10.7	28.6	21.4	14.3	29.8	17.9	38.1
VIC	15.3	27.2	28.7	24.6	26.6	26.4	29.4	21.8
WA	20.0	10.0	23.3	23.3	28.2	26.3	28.5	40.4

Note: The original percentages are based upon school average NAPLAN scores (plausible values). The modelled percentages are based upon the z-scores of the school-level residual. Modelled estimates are based upon replication of the ACARA modelling approach.

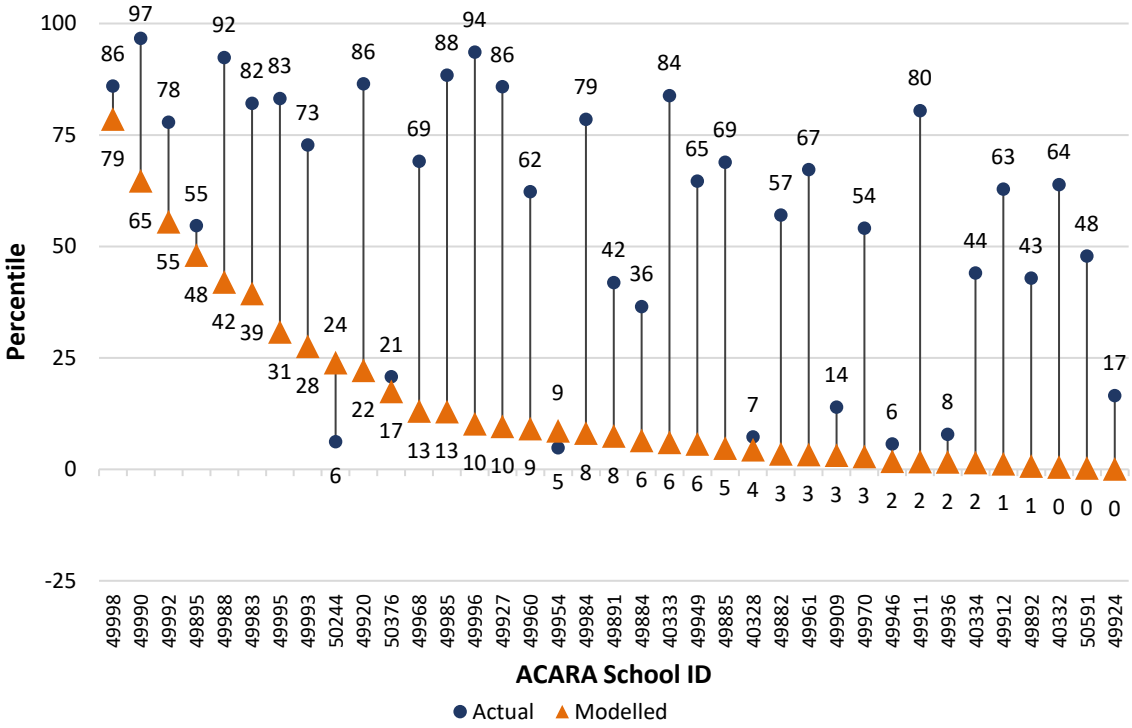
Percentile change following modelling: Numeracy

The change in the percentile position for individual ACT schools in Year 9 Numeracy is presented in Figure 2-1. The blue dots represent the school position using the original NAPLAN data; the orange triangles show the position of the school after modelling. The majority of schools fall over 50 percentage points, with one school falling from being in the 94th percentile

to being only in the 10th percentile of performers. The average change is a 41 percentage point reduction.

The movement in performance of ACT schools across the two measures is rather dramatic and somewhat unexpected. Many ACT schools move from being the top performing in the country, when looking at absolute scores, to being at the bottom, based on the modelled results. If the modelled results are to be believed then ACT schools, many of which are at the very top of the lists of schools in terms of students performing on average among the best in Australia, are actually performing among the worst in Australia, despite having the highest performing students. This is surprising. It is also surprising that this applies so systematically to all schools, both government and non-government.

Figure 2-1 Position change of ACT schools following modelling: Year 9 Numeracy 2019

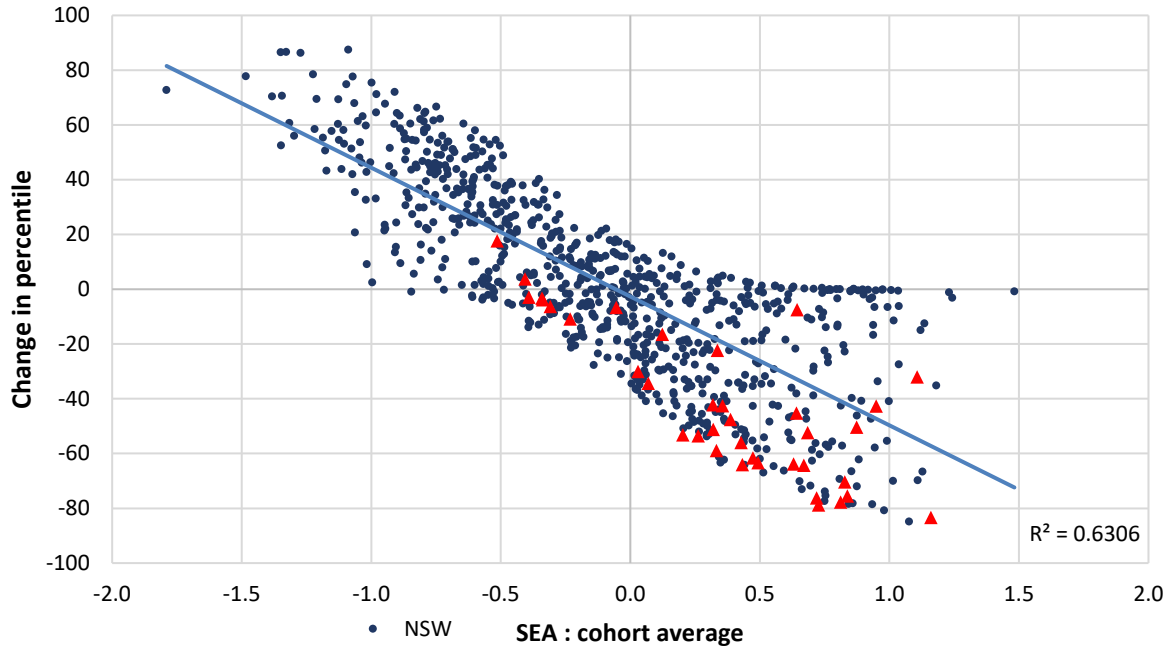


Source: Analysis of 2019 NAPLAN Year 9 Numeracy data. Modelled estimates are based upon replication of the ACARA modelling approach.

Analysis has also been undertaken of the change in percentile rank of New South Wales (NSW) and Victorian government and non-government schools based on the cohort average SEA scores of individual schools (see Figure 2-2 for NSW and Figure 2-3 for Victoria).

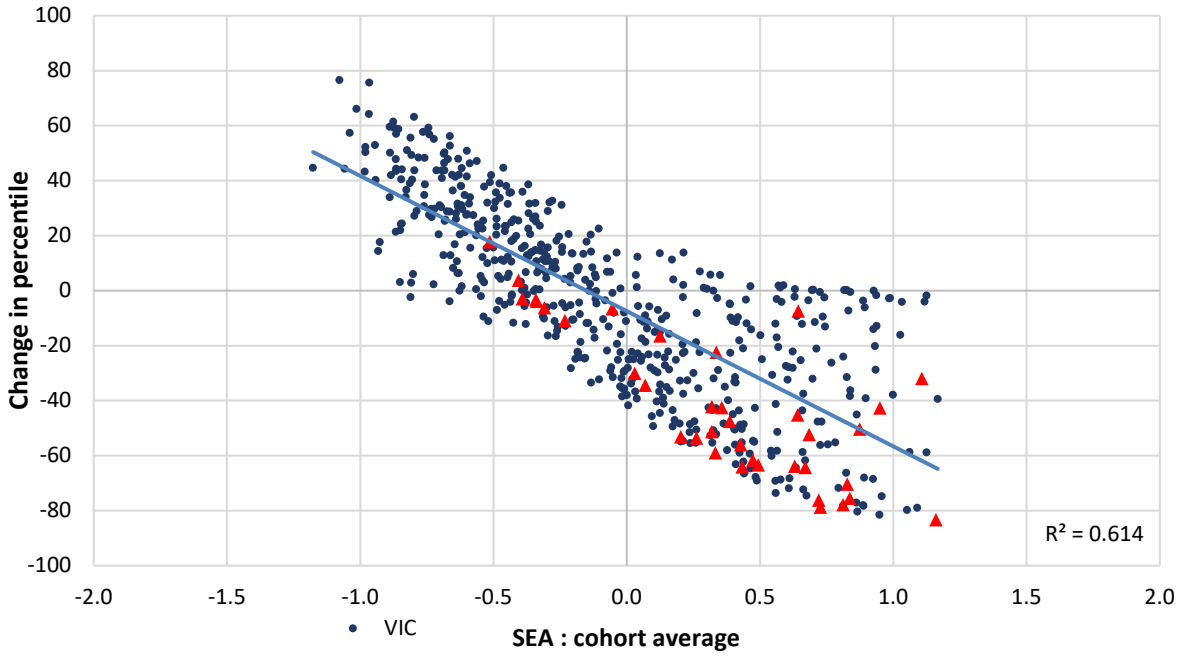
The results for ACT schools represented by red triangles are included for comparison. The changes are as profound in NSW and Victoria as in the ACT, across the SEA spectrum. The results highlight the negative relationship between school SEA (cohort average) and the percentile position change of individual schools. Lower SEA schools generally improve their percentile ranking, some by as much as 60 to 80 points. Higher SEA schools, by contrast, tend to fall in ranking, again some by as much as 60 to 80 points. The relationship between school SEA and change in percentile ranking is linear and fairly strong.

Figure 2-2 Position change of New South Wales and ACT schools: Year 9 Numeracy 2019



Source: Analysis of 2019 NAPLAN Year 9 Numeracy data. Modelled estimates are based upon replication of the ACARA modelling approach.
Note: School SEA is based on the Year 9 student cohort school average.

Figure 2-3 Position change of Victorian and ACT schools: Year 9 Numeracy 2019



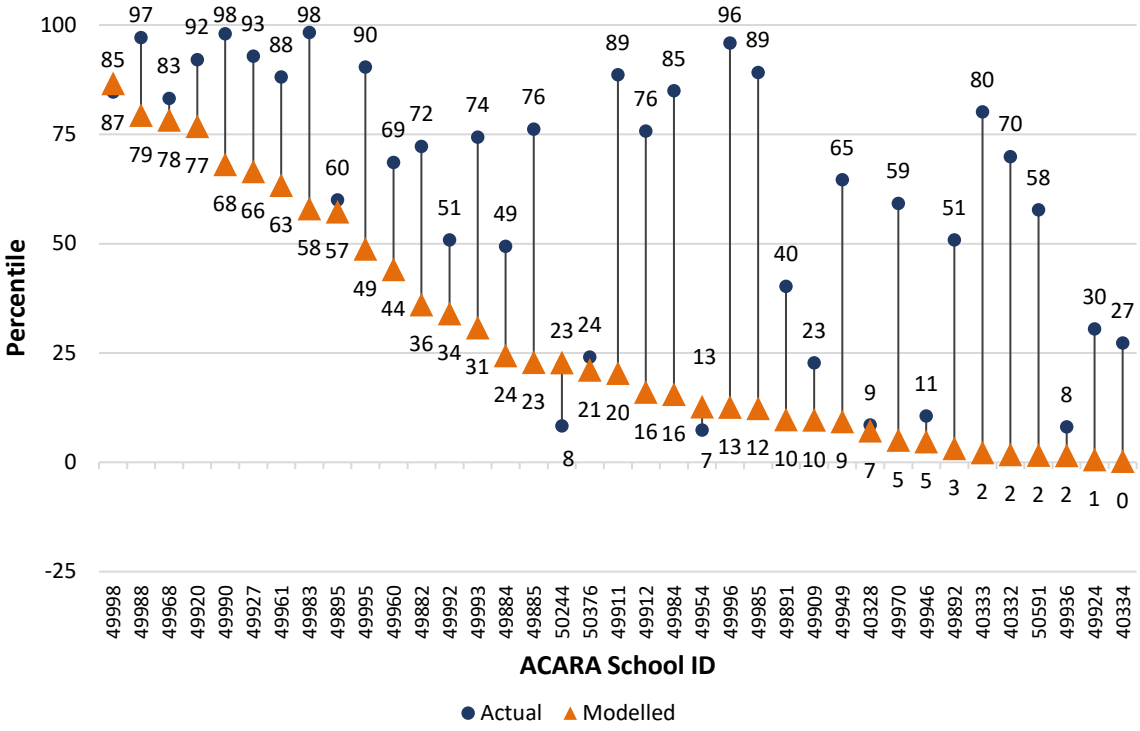
Source: Analysis of 2019 NAPLAN Year 9 Numeracy data. Modelled estimates are based upon replication of the ACARA modelling approach.
Note: School SEA is based on the Year 9 student cohort school average.

It should be noted for both Figure 2-2 and Figure 2-3 that the results include selective-entry schools. The selective-entry schools are almost all higher SEA, and because of selected intake and high-level average NAPLAN scores they change very little in percentile rank—they are the schools clustered towards zero change. The effects of excluding the selective entry schools from the analysis would likely be to strengthen even further the already strong relationship between school SEA and change in school performance percentile rank.

Percentile change following modelling: Reading

The movement in the percentile position of ACT schools in Year 9 Reading is less pronounced, with an average change of a 33-percentage-point fall in percentile rank (see Figure 2-4). There are also some very large movements, with one school falling from the 96th percentile to the 13th, and another from the 80th to the 2nd percentile.

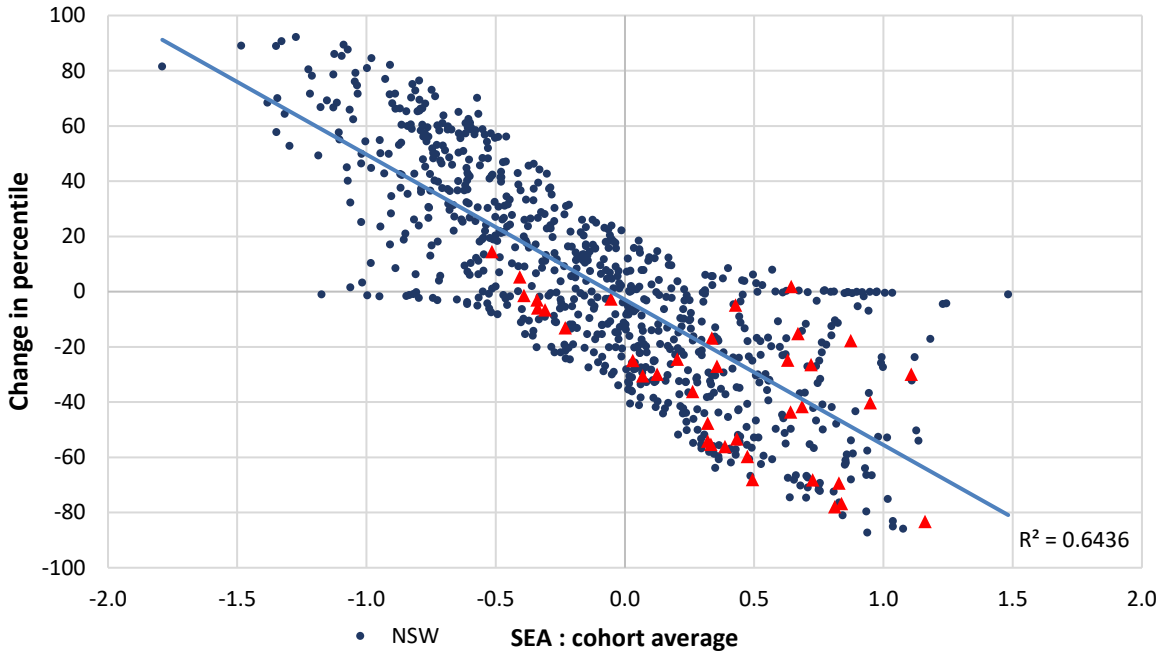
Figure 2-4 Position change of ACT schools following modelling: Year 9 Reading 2019



Source: Analysis of 2019 NAPLAN Year 9 Reading data. Modelled estimates are based upon replication of the ACARA modelling approach.
Note: School SEA is based on the Year 9 student cohort school average.

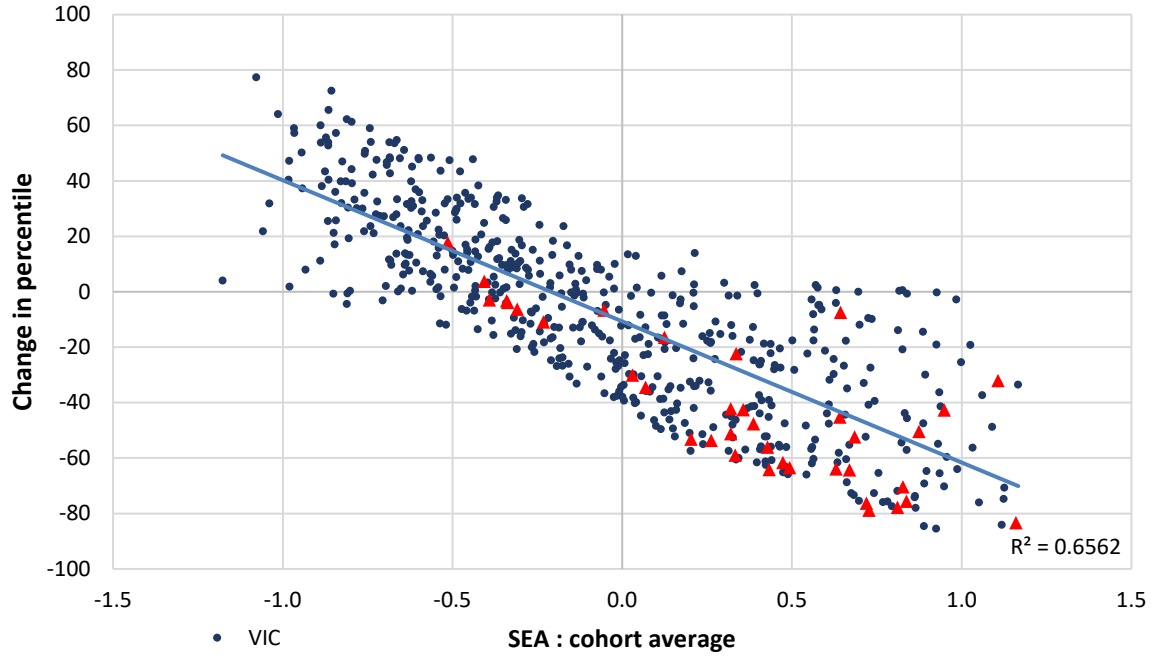
A similar position change in New South Wales and Victorian schools is also observed in Year 9 Reading (2019) to that observed in Numeracy (see Figure 2-5 for NSW and Figure 2-6 for Victoria).

Figure 2-5 Change in percentile rank of New South Wales and ACT schools, by school SEA: Year 9 Reading 2019



Source: Analysis of 2019 NAPLAN Year 9 Reading data. Modelled estimates are based upon replication of the ACARA modelling approach.
Note: School SEA is based on the Year 9 student cohort school average.

Figure 2-6 Change in percentile rank of Victorian and ACT schools, by school SEA: Year 9 Reading 2019



Source: Analysis of 2019 NAPLAN Year 9 Reading data. Modelled estimates are based upon replication of the ACARA modelling approach.
Note: School SEA is based on the Year 9 student cohort school average.

Student level analysis

Student-level analysis has also been conducted to see if the changes reported for schools are present at a student-level. Table 2-2 reports the quartile position of Year 9 students for Numeracy in each state and territory, using firstly the original (unadjusted) NAPLAN data and then using the modelled data. The results show that for students there is less movement between quartile rankings for students in the ACT. However, the proportion of ACT students in the top quartile does fall, from 28.2 percent to 17.5 percent. The movement in the proportion of students in the top quartile is much larger for the ACT than for other jurisdictions.

Table 2-2 Quartile position of students for Year 9 Numeracy: Percentage of students, by jurisdiction

	Quartile 1		Quartile 2		Quartile 3		Quartile 4	
	Original	Modelled	Original	Modelled	Original	Modelled	Original	Modelled
ACT	21.1	37.6	24.8	23.7	25.9	21.2	28.2	17.5
NSW	24.9	23.3	23.8	24.3	24.0	24.8	27.4	27.6
NT	44.4	28.5	23.8	26.6	19.7	26.0	12.2	19.0
QLD	28.0	27.4	25.9	25.9	24.9	24.8	21.2	21.9
SA	28.7	30.4	28.7	27.4	24.3	23.9	18.3	18.3
TAS	32.8	24.4	27.0	24.4	23.2	26.6	17.0	24.5
VIC	22.5	24.7	25.4	25.3	25.9	25.1	26.1	25.0
WA	20.9	20.8	23.6	23.8	27.0	26.7	28.5	28.7

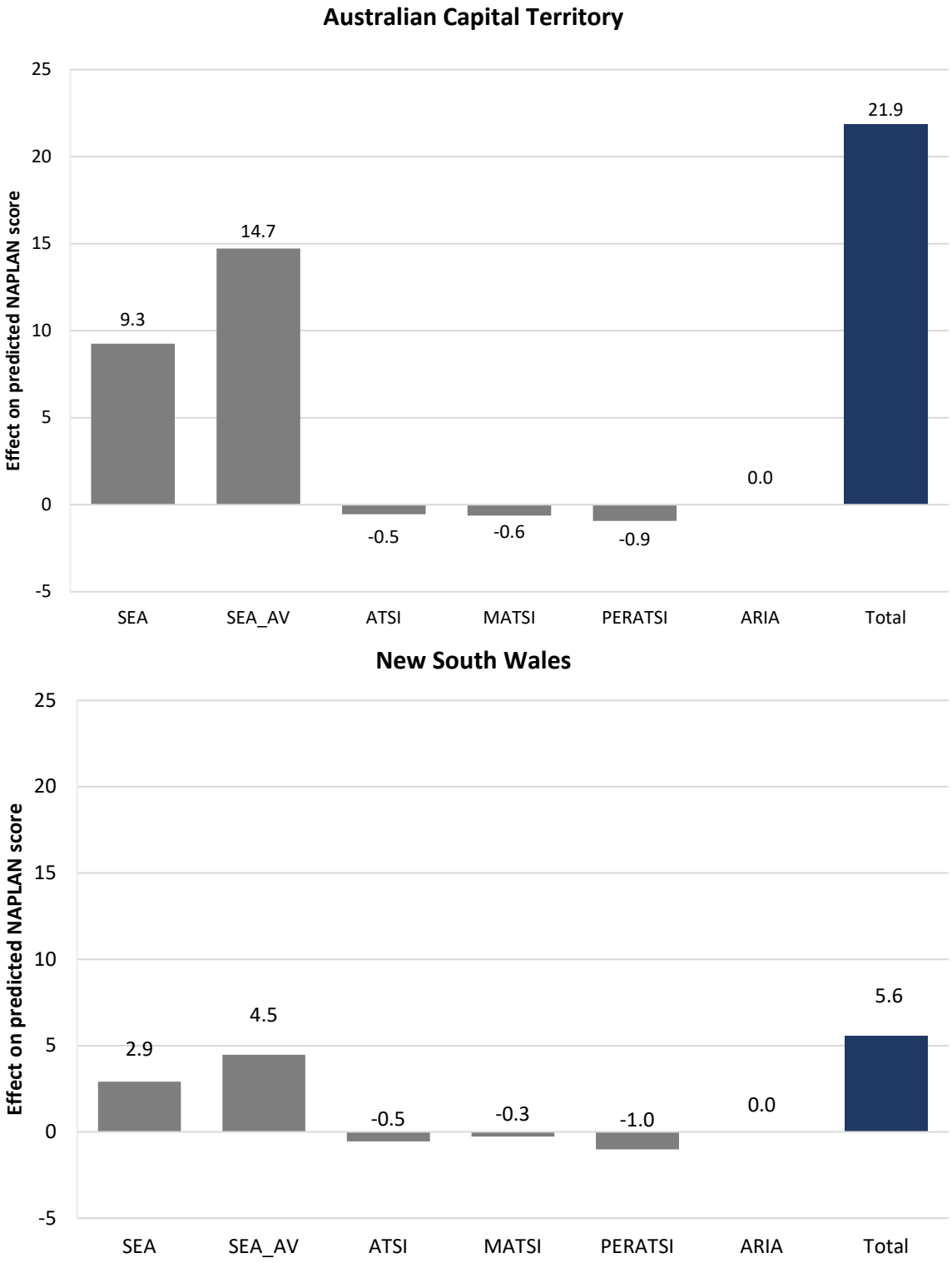
Note: The original percentages are based upon student NAPLAN scores (plausible values). The modelled percentages are based predicted student-level scores. Modelled estimates are based upon replication of the ACARA modelling approach.

Contribution of SEA to the change in the ACT position

The changes in the relative positions of ACT schools, and students, based on shifting from using the original (unadjusted) NAPLAN results to using the modelled NAPLAN results, are attributable to the inclusion of variables in the NAPLAN modelling to control for the characteristics of students and schools. In particular, the capturing of parent education levels and parent occupation types in SEA significantly alter the apparent position of ACT schools and students in relative terms.

The similar-schools modelling approach predicts the level of performance expected by students and schools. The sizes of the effects for the different variables used to control for various students and school characteristics, including SEA, are shown for the ACT and New South Wales (NSW) in Figure 2-7. After starting with a national average score of 596 in Year 9 Numeracy, the two SEA variables–student and year-level average SEA—push up the ACT expected performance by 24 points, compared to 7 points in NSW. SEA dominates the prediction of NAPLAN scores in the ACT, far more than for NSW and other jurisdictions (based on other results).

Figure 2-7 Contribution of student and school characteristics to the modelled results: Year 9 Numeracy 2019



Abbreviations: SEA = student-level SEA; SEA_AV=Year level SEA average within school; ATSI=Aboriginal and Torres Strait Islander status; MATSI=Missing Aboriginal and Torres Strait Islander status; PERATSI = Percentage of Aboriginal and Torres Strait Islander status students in school; ARIA= Accessibility Remoteness Index of Australia (ARIA) value of school location.

Note: Graph generated by using the jurisdiction level averages (student weighted) for each explanatory variable, multiplied by the coefficients from the Year 9 Numeracy similar-schools model. These averages were limited to schools where ARIA=0.

Between-school differences and SEA

Analysis was undertaken to examine the amounts of between-school variance in Year 9 Numeracy and Reading (2019) that the similar-school's model explains. Between-school variance is a measure of the variation in average NAPLAN scores attributable to schools as distinct from that attributable to the variation between students within schools. The greater the between-school variation, the higher the amount that schools contribute to student scores.

Table 2-3 reports the amounts of between and within-school variance estimated by the models. Six different models have been estimated, one using all schools and then separately for each of the five SEA quintiles. The reason for doing this is to see whether the between-school differences are much the same across the school SEA distribution or whether they vary in ways that relate to features of schools within each analysis, or possibly to the SEA measure itself.

Table 2-3 also reports the percentage of variance explained by each of the models. For the analysis involving all schools, between-school differences accounted for 30.3 per cent of total variance. Of this amount, the like-schools model accounted for 77 per cent (23.2% divided by 30.3%) meaning that 23 per cent was unexplained (7.1% divided by 23.2%).

Looking at the results for each quintile, it seems that for schools in quintiles 2, 3 and 4 there is little variation in student performance linked to between-school differences (8% or less). School differences in student achievement are much stronger in quintile 1 and quintile 5 schools. Focussing on Year 9 Numeracy in SEA_AV quintile 5, the between-school variance accounted for 22.3 per cent of total variance. Of this amount, the similar-schools model was able to account for 6.2 percentage points, meaning that the remaining 16.1 percentage points were left unexplained. For SEA_AV quintile 1 schools, between-school differences accounted for 20.4% of total variance. Of this amount the like-schools model explained 67% (13.7% divided by 20.4%).

Between-school variance as a percentage of total variance is much higher for SEA_AV quintiles 1 and 5, than for the middle three quintiles. The analysis shows that a much larger amount of the between-school variance is unexplained, denoted by the size of the grey columns in Figure 2-8. This is particularly the case in Year 9 Numeracy, where 16.1 percentage points of the 22.3 percent between-school variance share is unexplained.

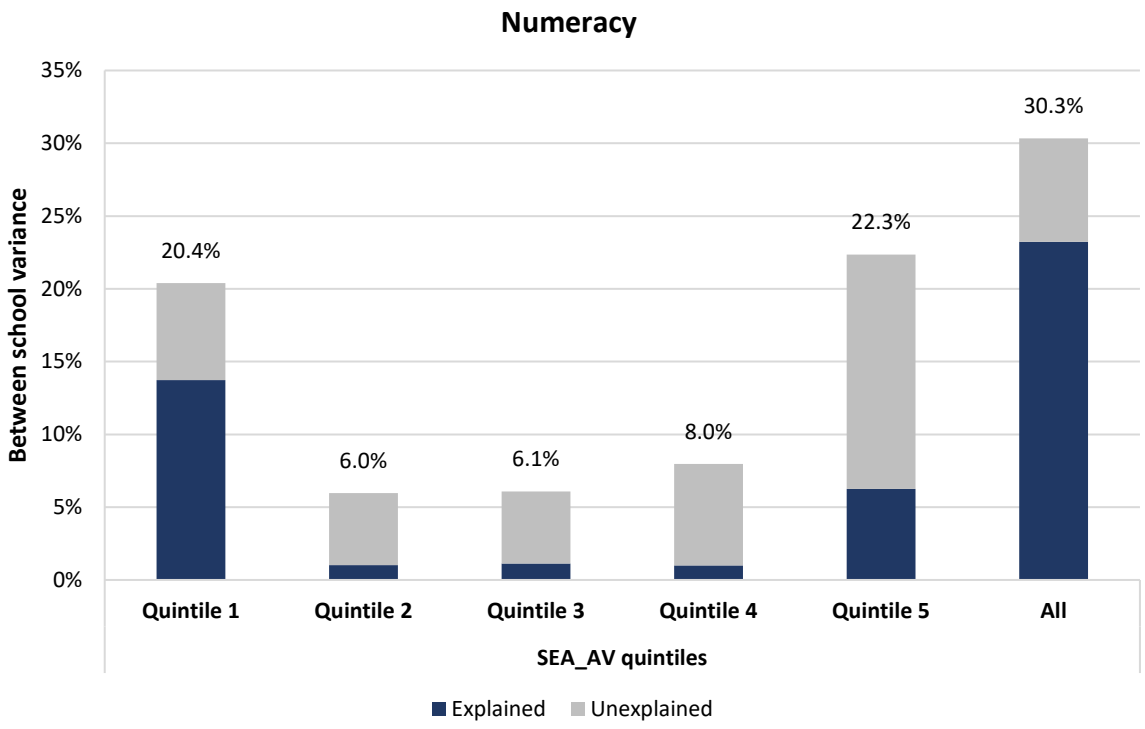
That the variance explained by the similar-schools model is much lower in quintile 5 schools than quintile 1 schools may mean that there are differences between high SEA schools that are not captured as well in the similar-schools model as those between low SEA schools. For example, most of the selective-entry schools in NSW and Victoria are in quintile 5 and their performance is not predicted as well by the existing like-school model controls. Also, it may be that the unexplained variance is also a result of the SEA measure itself not capturing as well the socio-educational differences across schools as well as it does for quintile 1 schools.

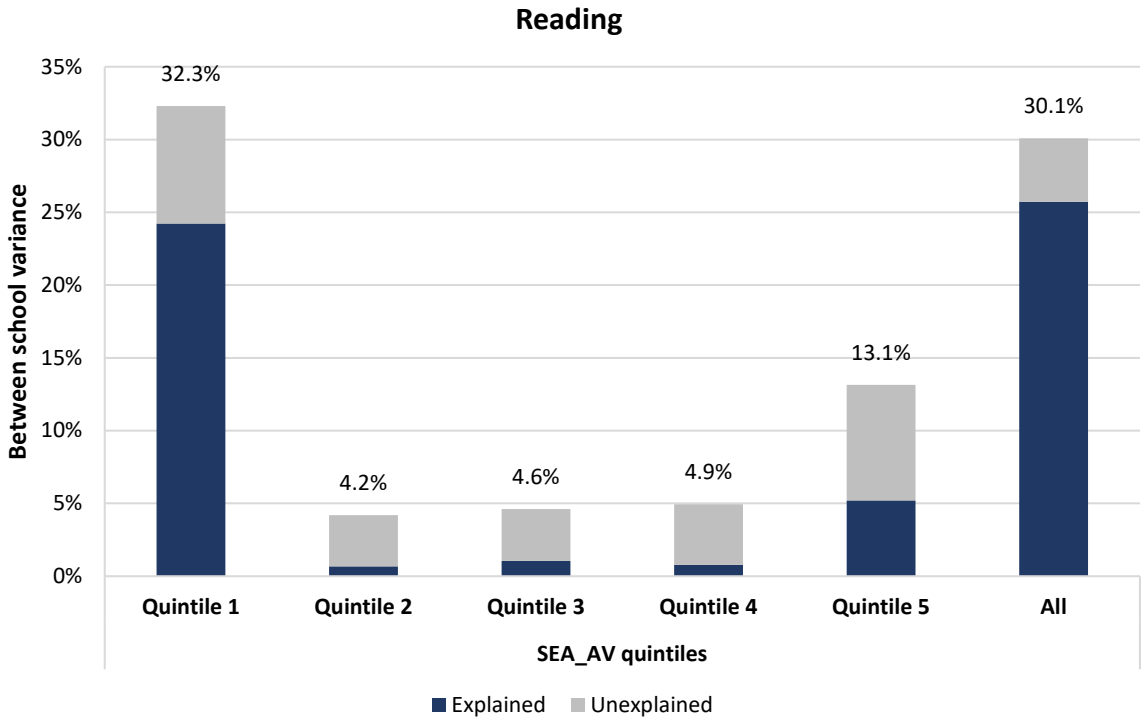
Table 2-3 Variance in Year 9 Numeracy and Reading (2019) and SEA_AV quintile

	SEA_AV quintile					
	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	All
Numeracy: Null model						
Between	669	176	183	238	823	1,213
Within	2,613	2,769	2,833	2,752	2,862	2,784
Total	3,283	2,945	3,016	2,990	3,685	3,997
Numeracy: Similar-schools model						
Between	472	133	126	131	248	201
Within	3,464	3,186	2,998	2,679	2,455	2,871
Total	3,935	3,318	3,123	2,809	2,702	3,071
Numeracy: Between-school variance						
Explained	13.7%	1.0%	1.1%	1.0%	6.2%	23.2%
Unexplained	6.7%	4.9%	5.0%	7.0%	16.1%	7.1%
Total	20.4%	6.0%	6.1%	8.0%	22.3%	30.3%
Reading: Null model						
Between	1,887	158	163	156	410	1,387
Within	3,956	3,627	3,370	2,997	2,711	3,225
Total	5,844	3,786	3,534	3,153	3,121	4,613
Reading: Similar-schools model						
Between	472	133	126	131	248	201
Within	3,464	3,186	2,998	2,679	2,455	2,871
Total	3,935	3,318	3,123	2,809	2,702	3,071
Reading: Between-school variance						
Explained	24.2%	0.7%	1.1%	0.8%	5.2%	25.7%
Unexplained	8.1%	3.5%	3.6%	4.2%	7.9%	4.4%
Total	32.3%	4.2%	4.6%	4.9%	13.1%	30.1%

Note: Variance estimates for each SEA_AV quintile were generated by running a 'null' model where the dependent variable was the total residual from either the original 'null' model or full similar-schools model.

Figure 2-8 Between-school variance in Year 9 Numeracy and Reading:2019 (%)





Alternative constructions of SEA don't change the results

Given the importance of SEA within the similar-schools model, further analysis identified the influence of how SEA is constructed. Alternative methods of constructing SEA-like measures were explored. Table 2-4 presents a description of the different SEA-like variables generated, and the methods used. This included using what are called the SEA Weighted Likelihood Estimates (WLE) generated by ACARA.

Other options included turning the reported parent/carer education levels into years of education, and using variables representing occupation type to estimate a single SEA score. Occupation type was used in two different ways:

1. a set of categorical variables, where each represented one of the four occupation types reported to ACARA
2. A numeric occupation status score for each of the four ACARA occupation categories as used in the Australian Socioeconomic Index 2006 (AUSEI06), which was developed based on education, occupation and income data in the ABS Census.[3]

Like in the estimation of SEA by ACARA, Senate weights were applied in approaches 2 to 4. Senate weights mean that each state/territory made an equal contribution to the estimation.

[3] The corresponding AUSIE06 values for the occupation categories are: Machine operators, hospitality staff, assistants, labourers and related workers = 27.9, Tradespeople, clerks and skilled office, sales and service staff = 39.6, Other business managers, arts/media/sportspersons and associate professionals = 60.9, Professional / Senior Manager = 83.1. Where individuals were not working, the AUSIE06 values were based upon educational attainment (McMillan, Beavis, & Jones, 2009).

Table 2-4 SEA construction approach options

No.	Name	Data approach	Method
1	SEA	Original values, as used by ACARA (5 plausible values)	Item Response Theory Generalised Partial Credit Model, with conditioning model
2	ALT_SEA	Single value (WLE) provided by ACARA, with provided plausible values if value missing	
3	SEA_PCA_SCORE	Turned parental education data into years of education, with a set of binary variables for occupation type.	Principal Components Analysis (PCA)
4	SEA_PCA_AUSEI06	Used years of education, and estimated occupation status values generated from the Australian Socioeconomic Index 2006 (AUSEI06).	
5	AUSEI06_SC	Australian Socioeconomic Index 2006 (AUSEI06)	
6	SFOE	Student Family Occupation and Education values, as used for Victorian government school equity funding	Weighted score constructed with pre-existing parameters.
7	SEAWLE	Using the WLE value provided by ACARA. But generated alternative set of plausible values.	Item Response Theory Generalised Partial Credit Model

Notes: Approaches 3-6: estimation at the parent level (within students), and student-level values by averaging parent estimates. Approaches 3-7: Multiple imputation using predictive mean matching used when missing data to generate five plausible values.

Sensitivity of the 'similar-schools' modelling to SEA construction

An extensive suite of modelling was undertaken to identify whether the reported performance estimates for ACT schools are sensitive to how the SEA value was constructed. This modelling applied an identical approach to that used by ACARA in the similar-schools modelling—the only difference being the specific SEA variable applied.

The modelling covered all year levels (3, 5, 7 and 9) and the Reading and Numeracy domains. Further, both the NAPLAN Weighted Likelihood Estimates (WLE) and PV score values were used.

Focussing upon Year 9 Numeracy, Table 2-5 reports the results for ACT schools. Schools are classified as performing above expected, as expected or below expected.[5] Under the ACARA modelling with the original SEA variable, 27 per cent of ACT schools performed as expected in Year 9 Numeracy. Modelling using the other SEA variable constructions did not result in

[5] This classification is based upon the z-value of the school-level residual.

significant change from this performance level, or the distribution of ACT schools across the three performance categories.

Table 2-5 also reports the proportion of the variance explained by the different models. This is akin to a measure of how well the model is able to predict the NAPLAN score. The original SEA model explains 76.5 per cent of between-school variance, and 8.9 per cent of within-school variance, and 29.4 per of overall variance. None of the models that included the alternative SEA variables could explain a higher proportion of variance.

These findings that alternative constructions of SEA do not materially change the results for the ACT should not be interpreted as meaning there are not anomalies in SEA and its construction.

Table 2-5 Variance explained and performance classification of ACT schools: SEA variables for 2019 Year 9 Numeracy PVs (%)

SEA variable	Variance explained			Performance classification		
	Between-school	Within-school	Total	Above	As expected	Below
SEA	76.5	8.9	29.4	0.0	27.0	73.0
ALT_SEA	73.5	7.2	27.4	2.7	21.6	75.7
SEA_PCA_SCORE	71.5	6.2	26.0	0.0	24.3	75.7
SEA_PCA_AUSEI06	73.1	6.6	26.8	0.0	24.3	75.7
AUSEI06_SC	69.3	4.3	24.0	2.7	24.3	73.0
SFOE	71.2	6.6	26.2	0.0	27.0	73.0
SEAWLE	73.2	7.1	27.2	0.0	21.6	78.4

Note: Classification of schools in the above table is based upon the z-score of the school-level residual.

Alternative applications of SEA in the modelling

The analysis also examined alternative ways of including SEA in the similar-schools model. The current similar-schools model includes the student-level SEA value, along with the year-level average within a school. Three different approaches to SEA were explored (see Table 2-6):

1. 'School' SEA average (in place of year level average), based upon the available data for a school in the NAPLAN dataset
2. SEA squared, at both the student and year level average level
3. proportion of students in top SEA quartile within a school, estimated using student-level data for Australia.

Table 2-6 Alternative applications of SEA in modelling

Variable	Rationale
1. 'School' SEA average (in place of year level average)	ACARA approach uses SEA average for year level. A 'school' average was generated using the available NAPLAN cohort data
2. SEA squared: Student and year level average	To test whether the effect of SEA on NAPLAN score changes as it increases. SEA and SEA_PCA_SCORE.
3. Proportion of students in top SEA quartile	Concentration of SEA Quartile 4 in a school is considered influential in explaining NAPLAN scores.

Note: Interactions of the above variables were also examined (e.g. [Proportion in top SEA quartile]×SEA).

The impact of alternative approaches to using SEA upon the similar-schools model were examined for Year 9 Reading and Numeracy. None of the alternatives resulted in an improvement in the estimated performance of ACT schools (see Table 2-7). However, these alternative approaches slightly improved the percentage of the between-school variance explained. For example, the model with the proportion of students in the top SEA quartile explained 77.8 per cent of between-school variance, compared to the original model explaining 76.5 per cent.

Table 2-7 Variance explained and performance classification of ACT schools: SEA application for 2019 Year 9 Numeracy PVs (%)

	Variance explained			Performance classification		
	Between-school	Within-school	Total	Above	As expected	Below
Original	76.5	8.9	29.4	0.0	27.0	73.0
School SEA average	77.0	8.9	29.6	0.0	18.9	81.1
SEA squared	77.7	8.9	29.8	0.0	21.6	78.4
Proportion in top SEA quartile	77.8	8.9	29.8	0.0	18.9	81.1

Note: The above results used the 'original' SEA plausible values. Results were not materially different when using the alternative 'SEA-like' variables.

Relevance of other variables and approaches to ACT results

The analysis in Figure 2-7 identified that SEA (at a student and year-level average) is dominating predictions in the ACT using the similar-schools model. Furthermore, the current similar-schools model explains 29.4 per cent of the total variance. This implies there may be other factors not included in the model that may contribute to differences in NAPLAN performance.

Several other variables and approaches were explored in the Year 9 similar-schools models (for Reading and Numeracy) to identify whether the results on ACT school performance are in any way influenced by variables currently not included as controls in the modelling. The other variables considered comprise (see Table 2-8):

- Selective-entry school status

- Language Background Other than English (LBOTE) status
- excluding schools with smaller cohorts.

Selective-entry school status was included as it is a proxy measure of a school being academically selective (testing and selecting students for enrolment on the basis of academic talent). In data provided by ACARA, there are there were 25 fully selective government secondary schools across Australia, most of which are in New South Wales. Only fully selective schools were identified in the modelling. It is important to note that there are many schools across Australia that are partially rather than fully selective and the selectiveness may influence ratings of school performance. In jurisdictions like New South Wales, Victoria, Western Australia and Queensland, for example, there are many partially selective secondary schools. It was not possible to identify partially selective schools for the current analysis. It is also important to note that there are no fully or partially selective schools in the ACT.[6]

LBOTE status was included as the ACT has a lower proportion of LBOTE students than Australia as a whole (ACT: 25%, Australia: 32%).

Finally, schools with smaller cohorts (Year 3 and 5 ≤10 students, Year 7 and 9 ≤30 students) were excluded to identify whether these schools were materially impacting the model results.

Table 2-8 Other variables and approaches

Variable or approach	Rationale
Variable: Selective-entry school status	A measure of whether schools have a 'high ability' intake. Selective-entry schools may be potentially biasing the results of non-selective schools.
Variable: LBOTE status	A possible measure of cultural influence on schooling., Language background may be an influence on student performance and may influence school performance where there are high concentrations of LBOTE students.
Approach: Omit smaller cohorts (Year 3 and 5 ≤10, Year 7 and 9 ≤30)	Smaller schools may be potentially biasing the model results.

Findings from modelling after including the additional variables are presented in Table 2-9. The inclusion of selective school status in the model leads to an improvement in the ACT results. The model featuring selective school status was able to explain 32.0 per cent of variance, compared to 29.4 per cent in the original similar-schools model. This improvement was attributable to between-school variance—the model containing selective school status explained 85 per cent of between-school variance, compared to 76.5 per cent in the original model.

[6] There are also a range of specific purpose schools in several jurisdictions, such as performing arts and science schools. These schools often fully select their students.

Table 2-9 Variance explained and performance classification of ACT schools using additional variables and approaches: 2019 Year 9 Numeracy PVs (%)

	Variance explained			Performance classification		
	Between-school	Within-school	Total	Above	As expected	Below
Original	76.5	8.9	29.4	0.0	27.0	73.0
Selective school status	85.0	8.9	32.0	5.4	24.3	70.3
LBOTE	77.4	9.1	29.8	0.0	24.3	75.7
Smaller cohorts	74.1	8.9	26.9	0.0	15.6	84.4

This analysis indicates that selective school status is an important variable in explaining differences in NAPLAN scores. Even though the variable only relates to 25 schools, it is sufficient to both improve the model's explanation of between-school variance, and the assessed performance of ACT schools.

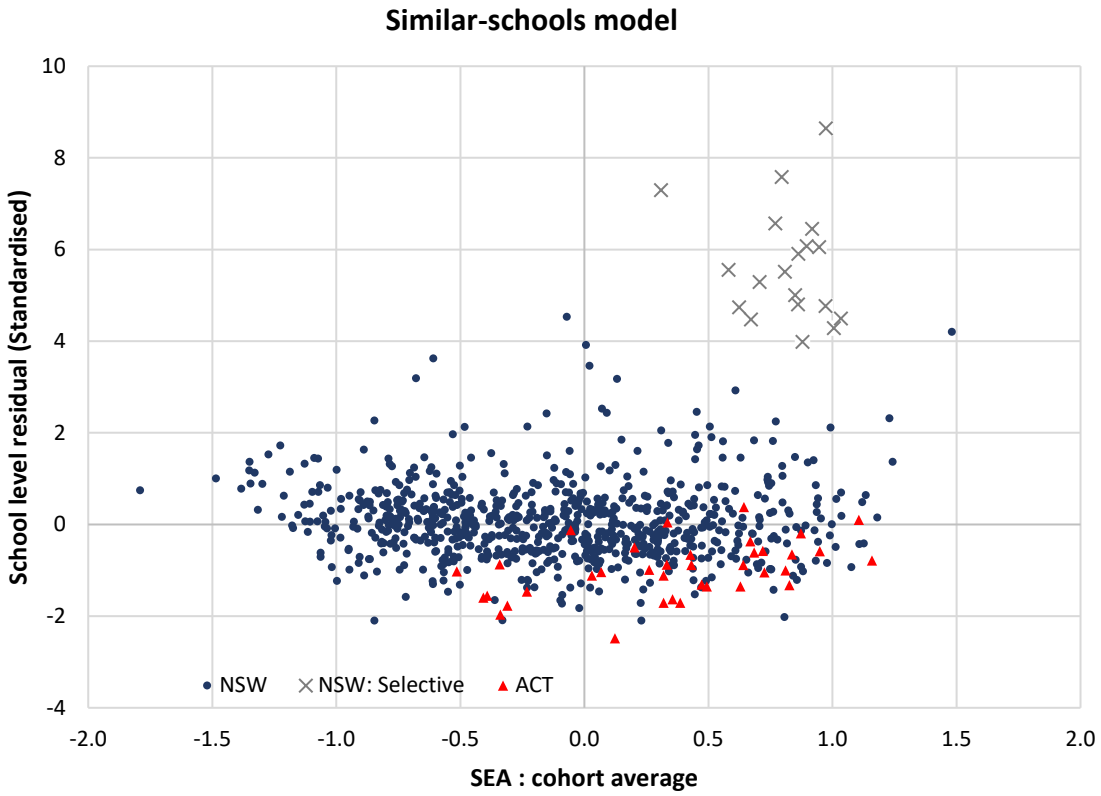
The importance of including a variable representing selective school status is visible in Figure 2-9, which presents for ACT and New South Wales schools the school-level residual scores (standardised) derived from the similar-schools model for Year 9 Numeracy. Residual scores are taken as a measure of school performance, with the figure separately identifying selective schools in New South Wales. Also reported in Figure 2-9 are the school-level residual scores (standardised) derived from a model which includes a variable denoting selective school status.

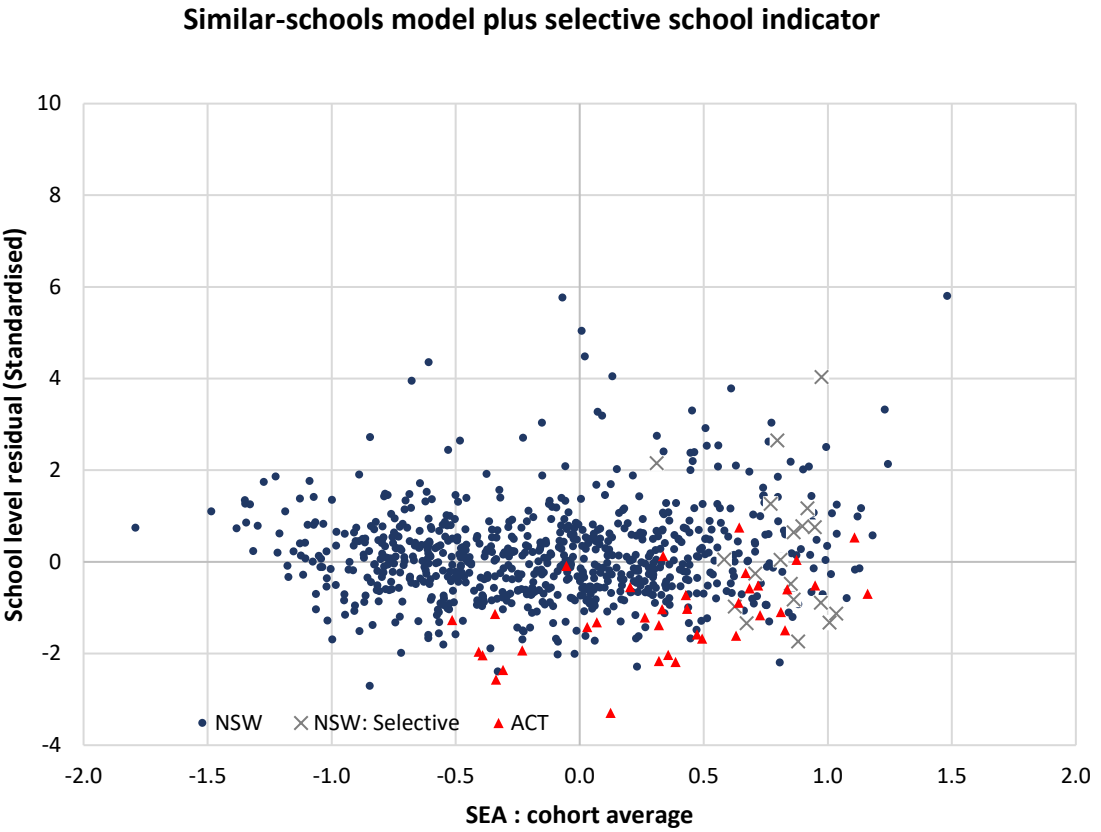
The introduction of a selective-entry school control has little effect on the performance position of ACT schools, however the effect upon selective -entry schools in New South Wales is very apparent. In the first graph in Figure 2-9 selective schools are predominantly in upper section of the upper right quadrant, denoting very high performance. In the second graph of Figure 2-9, selective-entry schools are spread among the other schools and no longer outliers.

An alternative to selective school status that could be explored in both Year 5 and 9 models is the proportion of students within a school in the top decile of performance two years previously (i.e. in Years 3 and 7 respectively). Such a variable would enable the model to take into account whether schools had an academically strong cohort.

The model using LBOTE status had a small improvement in the variance explained, whereas the model using smaller cohorts saw a reduction in variance explained.

Figure 2-9 School-level residual (standardised) and selective schools: ACT and New South Wales schools: Year 9 Numeracy (2019)





3. Summary of findings and options for next steps

Summary

The results of the analysis show that the similar-schools modelling leads to major change in the estimated apparent performance of ACT schools. When using the original or unadjusted NAPLAN scores, many ACT schools are high performing and among the top performers in the country. However, after the adjustments based on the application of results from the similar-schools modelling, the ratings of many ACT schools shift downwards and the schools are rated as being poor performers. This marked shift in performance assessment is attributable to the similar-schools modelling approach, and particularly the role of the SEA variables.

The results reported in this brief suggest that the quixotic results for ACT schools are unchanged by using alternative constructions of SEA, or by how SEA is applied in the similar-schools model. That does not mean that there are no anomalies in SEA and its construction. Rather, further work on how SEA is constructed and the way SEA affects the ACT, may be needed.

In terms of the similar-schools modelling, there is some evidence that missing variables may be important to study further. Selective school status, when included in the Year 9 Numeracy model, both improves the model explanation of variance, and the results for ACT schools.

There are several possible conclusions to the analysis findings with respect to the ACT, including:

1. the similar schools modelling approach, as currently applied, is not a suitable methodology for providing the basis of school performance reporting
2. there is a systemic NAPLAN underperformance issue common across ACT schools.

There is no evidence to suggest that there is a systemic issue in the ACT, given outcomes on other valued student outcome measures. For that reason, there are several options that might be explored as next steps to investigate alternative approaches to measuring school performance accurately and fairly.

Options for next steps

Following review of the analysis findings to date, there are several avenues that could be further examined, largely moving the analysis away from focusing upon SEA. Suggested options fall into three areas:

1. examining whether the poorer outcomes on the modelled school performance measures for ACT schools are reflective of real performance
2. refining the similar-schools modelling approach
3. exploring alternative approaches to analysing school performance.

Examining whether modelled outcomes for ACT schools are reflective of real performance

The first potential area for additional analysis involves closer analysis of high SEA schools in the ACT and other jurisdictions, examining performance across a range of different measures. The work would initially focus on schools that are performing strongly in the unadjusted NAPLAN scores, but not in the modelled scores. Comparison would be made with other schools also performing strongly in modelled data. The analysis would consider a wider range of performance measures and different sources of data. In the upper secondary years, this could include retention, transition and post-school outcomes.

Refining the similar-schools modelling approach

This additional analysis looks at further exploring the similar-schools modelling, including the role of additional factors not currently included. There is also potential value in documenting the range of assumptions underlying the modelling—both the modelling method itself, and its application for the similar-schools modelling.

A more technical area for examination is whether the truncated nature of NAPLAN scores, where there is an imposed score floor and ceiling, is having a greater impact upon ACT schools, than schools in other jurisdictions. A related consideration is whether the NAPLAN scores for the ACT are affected differently to other jurisdictions by the move to NAPLAN Online. In addition, the analysis reported above suggests that school selectiveness is an important factor influencing student outcomes. The analysis above looked at the role of selective school status. Selective school status is a partial measure of selectiveness. This issue could also be explored in models for years 5 and 9, by including the proportion of students in the top NAPLAN quintile within a school in years 3 and 5 respectively. Such an approach would control for the academic ability of school intakes.

With respect to the role of SEA itself, it may be worth examining whether variables representing the 'home learning environment' can provide a better explanation of student performance than SEA itself. Measures of the 'home learning environment', such as the number of books in a home, or whether a parent reads to their child, are not collected by ACARA. However, analysis comparable to the similar-schools modelling could be undertaken with the LSAC or LSAY. This analysis would seek to identify whether a model including a 'home learning environment' measure can further explain student performance, than one only using SEA itself. Other variables that may also be appropriate to examine include household income and family structure. Family structure, such as whether one of a child's parents is not employed by choice, may also be related to student NAPLAN performance. Finally, it may be appropriate to examine whether alternative information should be collected on parent/carer education and occupation. For example, adding an additional category to the non-school educational attainment variable to include Master's degree or above.

Exploring alternative approaches to analysing school performance

The third type or area of analysis would look at different ways of measuring school performance, using both NAPLAN data and other datasets. Three options with NAPLAN data

might be to utilise NAPLAN proficiency standards, adjusted learning gain or relative gain analysis.

The proficiency standard approach would seek to identify whether the proportion of students reaching a desired standard in a school is above, below, or the same as what would be expected given specific student and school characteristics. This option would also be underpinned by modelling. However, instead of modelling NAPLAN scores, the modelling would examine whether a student has reached the proficiency standard, or not. School-level actual and modelled average proportions could then be compared.

The adjusted or relative learning gain approach would use linked individual student data (e.g. results for Year 7 and 9), and analyse NAPLAN score gain. This analysis would initially look at the actual adjusted learning gain, and then move to use model-based methods to identify whether gain is also above, below, or the same as what would be expected. Use of a relative gain measure may reduce the need to include SEA and similar variables in modelling.

Other options in the third area comprise a time-series based analysis of PISA data, alongside a deeper analysis of ACT school performance. This analysis could include either a data-based exercise looking at other student outcomes such as ATAR and youth transitions. Another possibility is deeper engagement with ACT schools to identify what activities they are undertaking with a view to improving student outcomes. Both of these analyses would focus upon schools identified in the similar schools modelling as performing below expected, but identified using the original NAPLAN data as being strong performers.

Another option again within the third area is to undertake comparative analysis of approaches to reporting school performance. This would look at Australian and international jurisdictions, to identify lessons from how the performance of individual schools are reported.

References

Australian Curriculum, Assessment and Reporting Authority (ACARA). (2020). *Technical Report 2019: Approach to reporting on My School.* Sydney: ACARA.

Catholic Education Commission of Victoria. (2019). *My School does not make fair comparisons of school NAPLAN data.* Melbourne: Catholic Education Commission of Victoria. Retrieved from https://www.cecv.catholic.edu.au/getmedia/8c807fd5-783d-4f38-a01a-9978ac4c892e/CECVreport_MySchoolComparisonsUnfair.aspx?ext=.pdf

McMillan, J., Beavis, A., & Jones, F. L. (2009). The AUSEI06 A new socioeconomic index for Australia. *Journal of Sociology, 45*(2), 123-149.

NB A review of how SEA calculations are used in 'similar student' comparisons has identified anomalies for ACT schools. Investigations are continuing.